ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA  $\tt CCCACGCGTCCGGGCCGGAGCACGCCGCCGGGAGCTCCGGCTGCGTCTTCCCG$  $\tt CTGCTGCCGCCGCGGGGGGCCCAAGAAGCCGACGCCTGCCACCGGTGCCGGGGGGCT$ GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGAACA  $\tt CGGCTTGGGAGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATC$ CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA  ${\tt GCACCTGGAGGCCTGGAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT}$ GTGTGAAGACACTGAAAGTGTGCTGCTCCCAGGAACCTACGGTCCCGACTGTCTCGCATGC CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG  $\tt CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG$ GGCTGCACAGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA CGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAACCTGTGTGAGGAAAAACG  ${\tt GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT}$  $\texttt{GCCCTCCCGCGAAGACCTG} \underline{\textbf{TAA}} \texttt{TGTGCCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC}$  $\tt CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGGGGGAGAGGCTGC$  $\tt CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA$ AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG AGGAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCAGAAGC  ${\tt TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA}$  ${\tt TCACAAATTTCACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTC}$ ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG ATCTCAATTAGTCAGCAACCCAGTTTT

><subunit 1 of 1, 353 aa, 0 stop ><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKTLSKYESSEIRL LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESCKTCSGLTNRDCGECEVGWVLDE GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

#### Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343 and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and 313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGGATCCTCTAGAGATCCCTC GACCTCGACCCACGCGTCCGCCAGGCCGGGGGGGGGCGCCCAGCCGTCTAAACGGGAACA GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGCGCCCAGGTTGCGTAGGTGCG  ${\tt GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGC} {\tt ATG} {\tt GCCCGGAGGAGCGCCTTC}$ CGGGCCGCCGCAGGAGGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTAC CTGGCAAGCTGCAGGGCAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATA AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGT GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT TCTTTAAAACATGTCAACAAGCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT GAAAGACGCATCTGCGAGTGTCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG TACCCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTG GATTCTATGGAGTGAACTGTGACAAAGCAAACTGCTCAACCACCTGCTTTAATGGAGGGACC TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG CAAATGCCCACAACCCTGTCGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT GGAACCTGCCATGAACCCAACAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCCAGCTCAGGCAGC ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG **TGA**ACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA TGTGTTGAATGTTCAAATAATGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAG CATGATGGTATAGATTTCTTGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA TCAGGTTAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT GTTTGTTACATTTTTAAAAATTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACC AAACAATATAATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT AAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC CGCCATGCCCAACTTGTTTATTGCAGCTTATAATG

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094</pre>

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW

#### Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

CGGACGCGTGGGCGTCCGGCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG  ${\tt CCCCAGCCCACACCTTCACCAGGGGCCCAGGAGCCACC}$ TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG GAGCTAGCACCGGGTCTGCACCTGCGGGGCCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA GCAGGACCTGTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT  ${\tt TGCCTCGGCGTGCCACCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA}$ TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT  ${\tt GAACCACAGCGCCTTCTGGGGCATGACCCTGGA}{\tt TGA}{\tt GGGCATTCGCTACCGCCTGGGCACCA}$ TCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT CAATCCATTCTCTGGGACACATGACGCCTGTCCTGTCGCCCCAGAACCTGCTGTCTTGTGAC ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG  $\tt CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCC$ CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA  ${\tt TGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC}$ CTTGGGAGGCCAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC AGCTTCGTGCTGGGCCTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCATCACTGAGGCTG CGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGCCCCAATG  $\tt GGGCGGTGACCCCAGCCTCGCCCGACAGAGCCCGGGGCGCGCGGGGCGCCCAGGGCGCCTAAT$ CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGGAGCCGCGGGCAGACTGGCG ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC  $\tt CCCAATACCCCAATCCCGTATTCTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG$ TTGCCCAGGTTGGAGTGCAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACCACCTGGC  ${\tt TAATTTTGTATTTTGTAAAGAGGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTCGAACT}$ CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC TAAAACCAAAGTATTGATAAAAAAAA

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223</pre>

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDC ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPFPPIQGCMHGGRIYPVLGTYWDNCNR CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

#### N-glycosylation site.

amino acids 78-82, 161-165

#### Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300, 411-415

#### N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230, 269-275, 378-384, 442-448

#### Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site. amino acids 398-409

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTCGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGAATACCGCCGGCATGGGACCCACTCAG

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGCTACT CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG  ${\tt CAGCTTGAACTGCGTGG} \underline{\textbf{ATG}} \\ \texttt{ACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT}$ GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGGACCCGGCCAGCTATAGGCTCTGGGGGG CCCCGCTGCAGCCCACACTGGGTGTGGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG CTGCACCCCTGTCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT CCCTGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA  $\tt GGGGACAGGCACTCAGGAGGGCCCAG{\color{red}{\textbf{TAA}}} AGGCTGAGATGAAGTGGACTGAGTAGAACTGGA$ GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA GGGGCCAGGCCTCACATTCGTGGGGCCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT AATAAACACCTGTTGGATAAGCCAAAAAA

 ${\tt MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT} \\ {\tt LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHPDPPMALSRTPTRQISSSDT} \\ {\tt DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALRRAQ} \\ {\tt LGVVPQASVPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALRRAQ} \\ {\tt LGVVPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLRHPQASVPLCCCFHGPAFSTLNPVLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVPLRHPQASVP$ 

### Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

GCTTACAGCTGCTGATTCTCTGCTGTCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACC GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA CGTGCAGGTCACCGGGGGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGCTGAGAGTGAGAAG TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCAAAGACTG CGTGTTCACGGAGATCGTGCTGGAGAACAACTATACGGCCTTCCAGAACGCCCGGCACGAGG GCTGGTTCATGGCCTTCACGCGGCAGGGGCGCCCGCCAGGCTTCCCGCAGCCGCCAGAAC CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC  $\tt GGCGGCCCCAGCCCTCACG\underline{TAG}{\tt TCTGGGAGGCAGGGGGGCAGCCCCTGGGCCGCCTCCCC}$ GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCAGCTGGGAAGGGGCAGGCCGGTG CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTA AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCCACCCCCAACTCCCAGCCC CGGAATAAAACCATTTTCCTGC

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLIGKPSGKSKDCVFTEIVLE NNYTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT  ${\tt TTCCGTACTTCAGAA} \underline{{\tt ATG}} {\tt GGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT}$ GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC CTAGTGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC AATATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAA TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC CACCCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCACGACCCCCGGCCTG CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT TCCAAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTC CTGACTGGGATGGCAGAGAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTATC CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTCAGGAGCGCATAG TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA  $\tt CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGGCTTGATCGGGGGGCGCGGTGATATTT$ GTGCTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAAGGGGCGCTACACCTC CCAGAAGTGGAAATACAACCGGGGCCGGCGGAAAGATGATTATTGCGAGGCAGCCACCAAGA AGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAAATAACGAT CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC  $\mathtt{ACTGCCATACG}$   $\mathbf{TGA}$   $\mathbf{CAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA$ CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAAGTG CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELR VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN PWFCDCSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSCPTTTPGLPLFTP APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

#### Signal peptide:

amino acids 1-42

#### Transmembrane domain:

amino acids 542-561

#### N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

#### N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300, 522-528, 545-551, 633-639

#### Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

ACTTGGAGCAGCGGCGGCGGGGGAGACAGAGGCAGAGGCAGAAGCTGGGGCTCCGTCCTCGCCTCCCACGAGCG  ${\tt TCGCTTCCCAGGCGCCGGCGGCTGCAGCCTTGCCCCTCTTGCTCGCCTTGAAA} \underline{{\tt ATG}} {\tt GAAAAGATGCTCGCAGGCT}$ TGGTTTTCATCATTGACAGCTCTCGCAGTGTCAACACCCCATGACTATGCAAAGGTCAAGGAGTTCATCGTGGACA TCTTGCAATTCTTGGACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG GGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC AGAAGAAGTTGTGCACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG GCTCATACGTCTGCAGGTGCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAAAACCACGGATGTGAAC ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC TCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGAGTGACCATGGTTGTGAATACTCCT GTGTCAACATGGACAGATCCTTTGCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACGTGTG  ${\tt CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCGTGTGTAAGCAGTGAAGATTCGTTTGTGT}$ GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGGCTCG ATAATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT GCACTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTCG ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTTACCCAAG ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG CAGGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT  $\verb|CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT| \\$  $\tt CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGT$ TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC  ${\tt TGGAAAATCGCCTGAGATACAGA{\tt TGA}}{\tt AGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT$ GAACGCAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAAACAATCAGTACTGA GAAACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACACTAACTTGTATAAATTTATCTAGGAAAAAAATCCT TCAGAATTCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAATATACTGTGGACAC AACTTGCTTCTGCCTCATCCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT  $\tt CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG$ CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAAAATTCACCACTTCAG

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVFII DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV KRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGC EHECVNMEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI NEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCE HSCVSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA EDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA HMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN GITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDS PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247, 401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784, 781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500, 639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464, 540-546, 581-587

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG GCGCTGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGAGGTGTCTTC AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC ATGCCCTCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAC AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT GGGGCCCACCCTCAACCAATATCCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCACCTCTC TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCTGACTCCTTATGAAGCCAGCTG CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC CCCCTTGATCTGTACCCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC TGTTTGTATGAAAAA

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSKPAVQYQ WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKSSDTISKNGTL SSVTSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR MGAVPVMVPAQSQAGSLV

#### Signal peptide:

amino acids 1-29

#### Transmembrane domain:

amino acids 245-267

#### N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

#### N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262, 262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

 ${\tt CGCCACCACTGCGGCCACCGCCA}$ TTGAATTGTTCCTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCACAA TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA AGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT TAGATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAA GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA TATAATTACATATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC ACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGAT TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATA AGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGAT AATTCTGAAGAGGGGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCC ACCCACATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA GGTATAGGAGTCTATGTGCATTTTGGAATTACTCACCTGATACCATGAATGGCAGCTGGTCT TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT GACACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTC TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG TAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTTGCATGGATGTGC ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCA CAAGAATTTTTATATCTTTGGCTATCTAAGCCCAGCCGTGGTAGTTGGATTTTCGGCAGCAC TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACTTTATT TGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTTGGAGTCAT TAAGGTCTTGTGCAAGAGGGGCCCTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTT GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC TTTCCAGGGGATGTTCATTTTTTTATTCCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAAT  $\mathtt{ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGGTAA}\mathtt{ACATAGAGAATG}$ GTGGATAATTACAACTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA TGACTCATCAAATTATCCAATTATTAACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT GTTTATGCTATAGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT TTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT TGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTGCCTTTGAAACTAGTCC CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA TAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT 

MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL SNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT NSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLS SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF WNYSPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSIIAGL LHYFFLAAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGT TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA LALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV PCCFGCLR

#### Signal peptide:

amino acids 1-19

#### Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636, 648-664

#### N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181, 188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154, 155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329, 346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

#### N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394, 434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG CTAAGCGAGGCCTCCTCCCCGCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGA GGTGTGAGTGGGTGTGCGGGGGGGGGGGGGGGGTTGATGCAATCCCGATAAGAAATGCTCGGG TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG CCGCCGCGCCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC GGCACTCACAGCCCCGCAGCGCATCCCGGTCGCCCAGCCTCCCGCACCCCCATCGCCGG  ${\tt AGCTGCGCCGAGAGCCCCAGGGAGGTGCC}$ ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGCCCCCTCGCCTTCTCGGACGCGGGCCC CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC ACGGGCTCTCCAGCTGCTTCCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG CGTGCACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA CAGAGGCTTTCTTCCACTCTCATTTCCTGCCCATGCTCCCATGGTCCCAGAGGAGCCTG AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCTGGAGACCGACAGCATG GAGACCATGCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGGGCCCCCATTCTGCTCCCTCGA GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTCAGGAACAGGTGATCCACTCTGTA AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT CCCAGGCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC TTTTGTATATTAAAATGGAGTTTGTTTGT

MRSGCVVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD MFSSPLETDSMDPFGLVTGLEAVRSPSFEK

### Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA GAGCAGCCGGCTGCCGCGCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG CTGCTGCGCTACCTGGTCGCCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT  ${\tt CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA}$ CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATACACAATGAATACAAAAACTGGAACT CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT GCTCAGAGGAAAGGCTACTTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG  ${\tt CCGCGGCGGGCGGATCACGAGGTCAGGAGTTC}{{\tt TAG}}{\tt ACCAGTCTGGCCAATATGGTGAAACCC}$ CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC TTGGGAGACAGGAGATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC 

MARRSRHRLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE WKKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEED TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGIIAAVVVVA LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

#### Signal peptide:

amino acids 1-20

#### Transmembrane domain:

amino acids 130-144, 238-258

#### N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

#### Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

#### N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

#### Amidation site.

amino acids 226-230

AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT CATGCTGCTATTCCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA TTATATCATTAAGGAAATAGTAACCTTCTCTCTCTCCAATATGCATGACATTTTTGGACAATG CAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTCATCA  ${\tt TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAGCTAAG{\color{red} {\bf ATG}}{\bf AAGGACATGCCACT}}$ CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT ACTGCCTGAAAAATGTCTGTCCGAACTGAGCAACTTACAAGAACTCTATATTAATCACAACT AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCA ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT GGACTGGAAAACTTAGAAAGCATCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGT GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC TACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTG GATGAACATGAACAAAACCAACATTCGATTCATGGAGCCAGATTCACTGTTTTTGCGTGGACC CACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGATGGAAATTTGT CTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA GATATAAATGGCGTAACTCCCAAAGAAGGGGGGTTTATATACTTGTATAGCAACTAACCTAGT GCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTTGGTGTCCTGGAAAGCA AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAAACAGAAAAAAA TGTGTAAATGTCACCACCAAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA AGAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCC ${f T}$ **AA**AAACCACCAAGGAAACCTACTCCAAAAATGAAC

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT FPARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE ENKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDA LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL IKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAVDNLPD LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRC DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFRDMMEICLPLIAPESFPSNLNV EAGSYVSFHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE YEKNNTTTLMACLGGLLGIIGVICLISCLSPEMNCDGGHSYVRNYLQKPTFALGELYPPLIN LWEAGKEKSTSLKVKATVIGLPTNMS

#### Signal sequence:

amino acids 1-22

#### Transmembrane domain:

amino acids 633-650

#### N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583, 608-612, 624-628, 625-629

### Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

### Tyrosine kinase phosphorylation site.

amino acids 570-579

#### N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443, 491-497, 492-498, 634-640, 702-708

### Cell attachment sequence.

amino acids 277-280

GCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTTAG CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC  ${\tt CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGC} \underline{{\tt ATG}} {\tt AATCT}$ GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC CAACGACGCTGACCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG ACCTGATGATATTAGCACTGTGGTA**TAG**TGTCCAAACTGACTGTCATTGAGAAAGAAAGAAA TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT 

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNR
IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL
NAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK
ADEPDDISTVV

#### Signal sequence:

amino acids 1-33

#### Transmembrane domain:

amino acids 205-220

#### N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

#### N-myristoylation site.

amino acids 37-43, 45-51, 110-116

ACCGAGCCGAGCGGACCGAAGGCGCCCCGAG**ATG**CAGGTGAGCAAGAGGATGCTGGCGGG GGCGTGAGGAGCATGCCCAGCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT ACCGCGCTGTGCTGTCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCCATCCCCACCGAG ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCGCCAG CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCCTCAACCTCTCCTACAACCCCATCA GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCCACGCCCGAGTTTGTCCA GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGCG CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTTGTGGACGAGGGCCACACGGTGCAGTTT GTGTGCCGGGCCGATGCCGCCGCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT TCCATGCCCGCCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA GACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGGAGAGGCCAACAGCACCCGCGCCACTG TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC CTGGGCGTCGTCTTCTGCCTGGTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG ACGCGCCCCGCAAGTTCAACATGAAGATGATA**TGA**GGCCGGGGGGGGGGGGGGACCCCCG GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCCTTCCCACCTC CCAGCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGACGCGACACGCGGCAGAGTCA ATAATTCAATAAAAAGTTACGAACTTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT 

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCF VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFNLRTL GLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA FSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKVLEISH WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLILDSNPLA CDCRLLWVFRRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

### Signal sequence:

amino acids 1-41

### Transmembrane domain:

amino acids 556-578

### N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496, 505-509, 526-530, 542-546

### Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

### Tyrosine kinase phosphorylation site.

amino acids 590-598

### N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

 $\tt CCCACGCGTCCGCACCTCGGCCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC$ ATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCCGGGGATTCAGGCTCGCCAGCCCCAGCC  ${\tt AGGGAGCCGGCCGGAAGCGCG} \underline{{\tt ATG}} {\tt GGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCT}$ ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA CGAGGACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA  ${\tt GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC}$ ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT  ${\tt GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCATCATCACTGGTT}$ ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG ACAGGAAGATCCCAATGGTAAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCCAACTGCGATGATTAGGCCAGACCC  ${\tt TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC}$ GCCCTGATCTTCCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG  ${\tt CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCT}$  $\tt CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG$ TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG  ${\tt GGCAGTCAGGAGGGACGACAAGAAGGAATATTTCATC} {\color{red}{\bf TAG}} {\tt AGGCGCCTGCCCACTTCCTGC}$ GCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAA  $\tt CCGCAGGCCGCCTCCCGCTTGCTCCCCAGCCCACCCCCCTGTACAGAATGTCTGC$ CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAAACA AACAAAAAACA

MGAPAASLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVRTAKSLV TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ KLLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTHEAKGSDD APDADTAIINAEGGQSGGDDKKEYFI

Signal sequence:

amino acids 1-20

Transmembrane domain:

amino acids 331-352

N-glycosylation site.

amino acids 25-29, 290-294

Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304, 306-310, 334-340, 360-364, 385-389, 386-390

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GGGGGTTAGGGAAGGAATCCACCCCCACCCCCAAACCCTTTTCTTCTCCTTGG CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG  ${\tt TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAA} {\color{red} {\bf ATG}} {\tt CTGCTTTGGATTCTGTT}$ GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA TCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG AAAACAATGGCTTGCATGAAATCGTTCCGGGGGGCTTTTCTGGGGGCTGCAGCTGGTGAAAAGG CTGCACATCAACAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCCTTCCAGG ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACCTGCCAAC GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC CACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG CCCAAGAAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA GCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA TCCTCCCGGGCACTTTCAATGCCATGCCCAAACTGAGGATCCTCATTCTCAACAACAACCTG  $\tt CTGAGGTCCCTGCGTGGACGTGTTCGCTGGGGGTCTCGCTCTAAACTCAGCCTGCACAA$ CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTT TAGAAAGGATTTCATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT CGCCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT GTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAACCGAAAGC GACTCTTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTG  ${\tt TGGCTCTCACTCGCTCTCAGAC}$ CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA GCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG AACTGACAGTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCCTGTGGATTAG CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGCAGTTGCACGAAGGCATGAATGTAT TGTAAATAAGTAACTTTGACTTCTGAC

MLLWILLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH LFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQ TFLGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG NRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK IPGNWQIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSGLKMNCNNRNVSSLADLKP KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY LDTLSREKFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL VPGLLLVFVTSAFTVVGMLVFILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG AHRVYDCGSHSLSD

#### Signal sequence:

amino acids 1-15

#### Transmembrane domain:

amino acids 618-638

#### N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 122-126, 646-650

#### Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

#### N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 337-348

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGGGC TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCCTGTCGCGGCCTAAGGGAAACTGTTGGC CGCTGGGCCCGCGGGGGATTCTTGGCAGTTGGGGGGTCCGTCGGGAGCGAGGGCGGAGGGG AAGGGAGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC AGA**TG**AGGCCGGCGTTCGCCCTGTGCCTCTCTGGCAGGCGCTCTGGCCCGGGCCGG CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC AGCACCGTGCGTGCGGCCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG GCCCGGAGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGTTCCCACT GCACCCTGGAGAACGAGCCTTTGCGGGGGTTTCTCCTGGCTGTCCTCCGACCCCGGCGGTCTC GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCCAACGCTCCTGCACCGCGCGGAGATGCGC GGTACTCCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCCCGCGCCCCGGG GCCGCCTCTAACTTGAGCTATCGCGCGCCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCC GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGCCGCTCTTGTGTGACCA GCAACCAGCCCCGTGCCGCAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC ACCACTTGTCCCTGAACAAGACAATTCAGTAACATCTATTCCTGAGATTCCTCGATGGGGAT CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCT TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCA AGGAAGGAGTCTATGGGCCCGCCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC CAGTTCTGCACATTGCACAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG CAGAGGGTGCCTTGCTGGCGGAGTCCCCTCTTGGCTCTAGTGATGCATAGGGAAACAGGGGA CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC TTGTGTAACTGACAATTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAG GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA TTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTAT ATTGGTTCGAAATCCCAGGGAAAAAAATAAAAATAAAAATTAAAGGATTGTTGAT

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS TVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGA ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATA TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP SGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

### Signal sequence:

amino acids 1-16

### Transmembrane domain:

amino acids 399-418

#### N-glycosylation site.

amino acids 189-193, 381-385

### Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 98-102, 434-438

### Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

#### N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157, 185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469, 477-483

### Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

CGGACGCGTGGGATTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG CCCCGGCTACCTGGGCCGCCCCGCGGCGGTGCGCGTGAGAGGGAGCGCGCGGCAGCCGA  ${\tt CAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT}$ TATTGGCAGTGAAGGTTTTCCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG  $\tt CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA$ TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA  ${\tt AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTAAAAC}$  ${\tt CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC}$  ${\tt CGATATGATTATGTGGCTGTTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA}$  $\tt GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT$ TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA AAACTGCCTACAACTACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAACCATCACTCGCGATGGGAGTTTG CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTCAGCAGGCGGG CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG  ${\tt T}\underline{{\tt TAA}}{\tt CAGTGAACTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC}$ TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGG GACTGGTTGACTCTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT TGCCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA ATTTTAGAATTGAGTTGTGAAGATGTCAAAAAAAGATTTTAGAAGTGCAATATTTATAGT GTTATTTGTTTCACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA 

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTCGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAPIVSERNELLI
QFLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTTFPVTTGLKPTVALCQQKCRRTGTLEGN
YCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

### Signal sequence:

amino acids 1-23

#### N-glycosylation site.

amino acids 355-359

# Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

# Tyrosine kinase phosphorylation site.

amino acids 199-208

#### N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295, 305-311, 309-315, 320-326, 330-336

### Cell attachment sequence.

amino acids 149-152

CGGACGCGTGGGCGGCGCGCGCGCGCGCGCGGGCTGGGGCGTTCTT CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGCCCCCATGGCCCCCGAAGG GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGG CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAACCATCCGGGACAACTTTGGAG GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT GAGTGAGGAGCTGGTGGAGAGCTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCC AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCCTCCTGC CTTCCCTGTCCTGGGGGAACAGAGGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAGG GACACGAGGGGCAGCGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC TTTGGCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGG  $\tt CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT$ GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG  ${\tt ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC}$ AGAAGACGAGTTGGTGGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA CGCTGGCTGCTAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCCTGTGGCGGCCATG ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCCTTCATCAAGGGCAGA**TA** <u>A</u>TCGCGGCCACCACTGTAGGACCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG CCTTACAGAGCAGCCCAGGTACCCAGGCCCGGGCAGACAAGGCCCCTGGGGTAAAAAGTAGC CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA ATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCTGCTTAGGATTAGGTG GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT 

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFG IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

### Signal sequence:

amino acids 1-29

#### Transmembrane domain:

amino acids 372-395

### N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 290-294

### Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

#### N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179, 177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289, 326-332, 372-378, 395-401

### Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

# EGF-like domain cysteine pattern signature.

amino acids 181-193

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA GCACCATGCAGCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCTGGCCAGCCCCGGG GCCGCCTGACCGGGGAGCAGCTCCTGGGCAGCTGCTGCGGCAGCTGCAGCTCAAAGAGGT ACGTGGCCCTGCTGCAGCGAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTCAGCCAGAGC TTCCGAGAGGTGGCCGGCAGGTTCCTGGCGTTGGAGGCCAGCACACCCTGCTGGTGTTCGG CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGCGCGCTGTCCCCGCGCAGCGCCCGGGCC CTCCAGGCTGGTGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA ACTTCTGGCAGCAGCTGAGCCGGCCCGGCAGCCGCTGCTACAGGTGTCGGTGCAGAGG GAGCATCTGGGCCCGCTGGCGTCCGGCGCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT CCTGGCTTATGAGTGTGGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC ATCAAGGAGGAGGCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA  $\tt GTGCAGCTGTGCCTCGGATGGTGCCCTAGTGCCAAGGAGGCTCCAGCCA{\color{red}{\textbf{TAG}}} GCGCCTAGTG$ TAGCCATCGAGGGACTTGACTTGTGTGTTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG GCGATGACTGCATGGACAAATGCTCTGTGCTCTAGTGAGCCCTGAATTTGCTT CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTCATTGTTTACTTGTCCTGTCAC TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT AAAA

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

### Signal sequence:

amino acids 1-18

### N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

#### Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

 $\tt GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCG{\color{red} \underline{ATG}} GGGACAA{\color{blue} A} \\$ GGCGCAAGTCGAGAGGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAAGTTTGACCA AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG ACCTCGTCATTCCTGGTG**TGA**GCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC TACACCCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC ATCCTCCTTCATGCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAA GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC 

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW
KFDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

### Signal sequence:

amino acids 1-27

#### Transmembrane domain:

amino acids 238-255

### N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 270-274

# Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158, 193-197, 203-207, 287-291

#### N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

 ${\tt CAGCGCGTGGCCGCTGTGGGGACAGC} \underline{\textbf{ATG}} \\ \texttt{AGCGGCGGTTGGATGGCGCAGGTTGGA} \\$ GCGTGGCGAACAGGGGCTCTGGGCCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCT GGAGGCCGCCGGGCCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG GCTCGTGCCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGG  $\tt CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGGATGAGGAGGAGTGCAGGATTGAGCC$ ATGTACCCAGAAAGGGCAATGCCCACCGCCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTAGCA GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCC ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTC CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC CTCCGCCCACTGGGGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC  $\tt CTCGCTGCCC{\color{blue}{\textbf{TGA}}} GGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA$ GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT  $\tt CTGGCCACGTGGAACCTGGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC$ TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAG GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR TSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLL LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

### Signal sequence:

amino acids 1-30

# Transmembrane domain:

amino acids 230-246

### N-glycosylation site.

amino acids 126-130, 195-199, 213-217

# Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

### N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218, 224-230, 230-236, 263-269

# Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

# Leucine zipper pattern.

amino acids 17-39

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCAGCAGAGGTCGCGCACAGATGCGGCTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAGCACCTCAGTCCCCCCAGAGACTCTTG GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGTT CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA TAATTCCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG  ${\tt CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCT}$ GCCACCGCGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTC TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTCGTCATC CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGCCTCCCCGGAG TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG ACGAAGCTGTGAGTGCCGTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG GGCTGCCCTTACCCGTGGACGACCAGAGCCCCCCAGCATACCCCGGCTCAGGGGACACGGA CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA GTTGTTCCTAAGAAAC**TGA**TTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTCTTGGTTTTAGACAAATGTAAACAA AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC CTGTTTCTTGACACAGACTGATTAAAAATTAAAAGNAAAAA

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT WPSTHETLLTTWKIVAFTATSVLLVLLLVILARMFQTKFKAHFPPRGPPRSSSSDPDFVVVD GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGDTDTGPGESETCDS VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEEVASTSPGIHHAHWVLFLRN

#### Signal sequence:

amino acids 1-41

#### Transmembrane domain:

amino acids 325-344

#### N-glycosylation site.

amino acids 104-108, 134-138, 192-196

### Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366, 364-368, 380-384, 467-471, 468-472

### N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424, 478-484

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 394-405

 ${\tt TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCGCTCCACCGCCGT}$  ${\tt TCGGATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGCACAGAGGCCTTGTTA}$ TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT  ${\tt CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG}$ TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCCTACATGTTCCAGTGGAATGATGA  $\tt CCGGTGCAACATGAAGAACAATTTCATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT$ CTAGAGAAGCTGAAGGTGAGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT  $\tt CCTAATCCCCAGCATTCCCCTTCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT$ GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACCATC TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA  ${\tt AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCATTCCGAGTGTGTT}$  $\tt CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA$ GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG GTTAT $\underline{\mathbf{TAG}}$ GACATATAAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC CTCTTATTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT  $\tt CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA$  ${\tt GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG}$ GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACCTCTGTTTTCCTTGCTCTATACAG CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTCAAAGAATAAAATCAAATAAAGA GCAGGAAAAAAAA

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRREEKQSNSTACQDL
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLV
VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

#### Signal sequence:

amino acids 1-21

#### Transmembrane domain:

amino acids 235-254

#### N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

### Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226, 299-303, 306-310, 323-327

#### N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

ATCCGCAGGTTCCCGCGGACTTGGGGGCCCCCCTGAGCCCCGCGCGCCCCCAGAAGACTTGT  $\tt GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCCAGCGAGGGCCTACCACC \underline{\textbf{ATG}} \texttt{ATCACTGGTGT}$ AGCGGCGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTC AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA GGCAGAGACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT GGACAACGTGGCTGCCGAGCAGCACCACCACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG CACGGATGATCGAACAGAGGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC CATGGACTCTGCCACTGCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTTGACCACAAATGGCCACCGTTTGCT GTTGACCTGACCATGGAACTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA  ${ t ACTCAGGTGATGGAAGTTGAAGAGGGT}{ t TAAAAAGCAGGATGTGTTGATT}$ TTAAAATAAAGTGCCTTTATACAATG

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGAR SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE GPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMGIDSSDKVD FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFLHILES NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQVMEVGNEE

### Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 218-222

Casein kinase II phosphorylation site. amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site. amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature. amino acids 50-65

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT  ${\tt TAAATTTCAGCTCATCACCTTCACCTGCCTTGGTC{\color{blue} ATG}CCTTGCTATTCTCCTTGATCCTT}$ GCCATTTGCACCAGACCTGGATTCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT  ${\tt CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG}$ ACGCTGGGACATTAAGGACGTGGCTGTTGTTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAAGGTCCTCATCCA ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG ATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA GTCCCAGAGGGTGTCAGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGG TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAG  ${\tt TGTGGCTGTCATCTGGTG} \underline{{\tt TAG}} {\tt GTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA}$ GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCCTGATTCTCAGGCCTTCAGAGTTGG ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG TATGTGTTCAAA

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLGCGRAVL TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

### Signal sequence:

amino acids 1-15

### Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238, 267-271, 294-298, 316-320, 336-340

### N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143, 180-186, 263-269, 286-292

#### Amidation site.

amino acids 196-200

Speract receptor repeated domain signature.

amino acids 29-67, 249-287

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGGATCCTCTAGAGATCCCTCGACCTCGAC CCACGCGTCCGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCCTGCTGTTCGGCTGCCTGGG  $\tt CGTCTTCGGCCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG$  ${\tt TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT}$ GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC  ${\tt TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC}$ GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA TGTGGACAAGAGGGTCATGGAGACAAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC TGCCCTCCATGATCAAGAGGGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG  $\tt CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCCAGAAAAGAGCGGAAATCC$ AAGAACTCC**TAG**TACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSIQGKMSI PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

#### Signal sequence:

amino acids 1-21

#### Transmembrane domain:

amino acids 104-120, 278-292

#### N-glycosylation site.

amino acids 228-232

# Glycosaminoglycan attachment site.

amino acids 47-51

# Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

### Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

#### N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

#### Amidation site.

amino acids 265-269

# Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

CCCACGCGTCCGCTGTTTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA AAAAAAAAACACCAAACGCTCGCAGCCACAAAAGGG**ATG**AAATTTCTTCTGGACATCCTC CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCCTAA GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTCATACCTTTGT GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA GGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG TTTCATAAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG TCTGTGTCCTAATTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA CCTGGCAGTTTTAAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA  $\texttt{AAGCGCAA} \underline{\textbf{TAA}} \texttt{GCACCTAGTTTTCTGAAAACTGATTTACCAGGTTTAGGTTGATGTCATCTA}$ ATAGTGCCAGAATTTTAATGTTTGAACTTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA TCATTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA TGAAGGACTATATCTAGTGGTATTTCACAATGAATATCATGAACTCTCAATGGGTAGGTTTC ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG 

MKFLLDILLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

### Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 30-34, 283-287

Casein kinase II phosphorylation site.

amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

TTTCGCTGGTCCTGTTGATGCCTGGCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG AGGTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT AGCACGGGATTTATACAGTGCACTAATTCAGTTTTTCCAGATATTTCCTGAATATAAAAATA ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA TGGATATTCTGATCCCGAATCAATTATAGGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC AGGAAGCAGAACTGGTTTGAGGCCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACTATAGTTGAAAAAGTACTTGCGAGA AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCCTTTTGACATGATTAATCGA  ${\tt TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGA} {\tt TAA} {\tt ACTACCTTCCCAAAAGAGAACAT}$ CAGAGGTTTTCATTGCTGAAAAGAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAA TTATCTTTCATATCTGCAAGATTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC TTTTGTTTTTGGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA AAAATATTATATATAAAAGTAAAAAAAAAA

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL VGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH GPYVVTSNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT EIMNNYKVLIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKAEKKVWKIFKSDSEVAGYIRQ AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

### Signal sequence:

amino acids 1-22

# N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

# Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352, 353-357, 424-428

# Tyrosine kinase phosphorylation site.

amino acids 423-432

### N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

# Serine carboxypeptidases, serine active site.

amino acids 200-208

# Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT  ${\tt TTTTCCCTTACAGGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT}$ CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGAAG TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC CTGGGGGAGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA GACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGGGAGAAAGTATGTTAAAAATA GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGG**ATG**GGGACCCTGGGTC AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTTGGTCTGACCACTCTGCCTTGTGTTTT  ${\tt CCTCACCGCCCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG}$ ACAGCCGCTCTGTGGTCCTCCTCAGTGGTCTGGGTGCTGCTGGCCCCCCAGCAGCCGGC ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT CCACCAAGGGACGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG CCCCTCATCGTGCAGCCCTGCAGCGAAGTGCTCACCCTCACCAACAATGTCAACAAGCTGCT CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT TTTGTCTCCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT CTTCTACATCTACGGCTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG CAGTATCACCACCCGCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT  ${\tt GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC}$ TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC ATTCACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG GCAACTTTATTTTCTTGGGGAACAAAGG**TGA**AATGGGGAGGTAAGAAGGGGTTAATTTTGTG ACTTAGCTTCTAGCTACTTCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA TTTCAATATTTCCCAAACTTTAAGAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV YKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLACGSL YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY FPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGN LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG YSVVFVGTKSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

### Signal sequence:

amino acids 1-32

#### Transmembrane domain:

amino acids 71-87

### N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

# Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387, 384-388, 471-475, 481-485, 530-534

### N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

AGGCTCCCGCGCGCGCTGAGTGCGGACTGGAGTGGGAACCCGGGTCCCCGCGCTTAGAGAACACGCGATGACCA GCAGGCTGGACCCCTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCCTGATGGCCGCAGAGATCG GTTCCTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACTGC TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT  ${\tt ACTGGACGGGGTGGTTTGACTCGTGGGGGGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT}$ CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCT TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCGTCGGCCACGTGCATGATCGGGGGCAGG  ${\tt TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT}$ ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG GCTTAATTGGAAATCTCTATCTGAATGATTCACCCCTGAAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA GCTTCTTTCAGAGGTTCGGCCTGGACAAATGGNGTTCCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTTGG GTAGCTTGTCCATCAGCTCCACGCCTTGTGACACCTTTCTGAAGCTGGAGGGGCTGGGAGAAGGGGGTTGTATTCA GCAGCGGAATCAACCAGGTCATCGTTTTTGAGGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCCC  ${\tt ACCTGGGCAGGAACCAGTACATTAAG{\color{red}{TGA}}{\tt GCGGTGGCACCCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC}}$ CTCTTGACCTGAAGCCTGGTGGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGG ACTGGGGGCTACAGTCTGCCCCTGTCTCAGCTCAAAACCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC TGGCTTTGTTGATGATGGCTTTCCTACAGCCCTGCTCTTGTGCCGAGGCTGTCGGGCTGTCTCTAGGGTGGGAGC AGCTAATCAGATCGCCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGCACCGGACGTCACAGCCC  ${\tt TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT}$  ${\tt TTTATCCCCGAAATCCTGGGTGTCACCAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT}$ CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCA AGGGAGGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA GGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG ACAGAAGGCCCAGCTCAGTGGCCCCCGCTCCCCCACCCCCGAACAGCAGGGGCAGAGCAGCCCTCCTTC GTTGCAGTAAAGCTATAACCTTGAATCACAA

MTTWSLRRPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW IFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKFDFSGNLDLEAFVLMA AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ YKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKDGLSKGIVQGVLAT INLQSTHELQLLTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILDSSEVLKTVSAIVDAGS SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE TSITSSGILSGHVHDRGQVFVNTVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN IDDQRKGLIGNLYLNDSPLKNFRIYSLDMKKSFFQRFGLDKWXSLPETPTLPAFFLGSLSIS STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA LQFTETPHLGRNQYIK

### Signal sequence:

amino acids 1-27

# Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

# N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315, 320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCA**TG**GCTCCC AAGAAGCTGTCCTGCCTGCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA GGCAGACACTCGGTCGTTCGTAGTGGATAGGGGTCATGACCGGTTTCTCCTAGACGGGGCCC CGTTCCGCTATGTCTCGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC CGGCTTTTGAAGATGCGATGGAGCGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT CTGGGGCCAGAATCACTCCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC TATATCTGAAGCAGGGGACCCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCGAGCCCCAAGATGATGCTTGGACCTGTG ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC GAACCTATATGACCCATACCATTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC CATGACCGTGCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG AGACAAACTATTTTTGACGGGGAAACTGGGTCCAAACTGGATATCTTGGTGGAGAACATGG GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG ACCAAGGCCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTTTTCCTAGGGGAGCCCTCAACAAAATTA CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTTGGATAAGCCTATC CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG  ${\tt TGCCTCTGAACCAATGGAGTTAAGTGGGCAC}{\tt TGA}{\tt AAGGTAGGCCGGGCATGGTGGCTCATGC}$ CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA GACACTCCATCTCAAAAAAAAAAAA

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMML
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYL
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

#### Signal sequence:

amino acids 1-27

#### N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 4-8

#### Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554, 603-607, 644-648

#### Tyrosine kinase phosphorylation site.

amino acids 191-198

### N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233, 231-237, 274-280, 296-300, 307-313, 447-453, 484-490

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC ACCCACAAT**ATG**GCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT CTTTCGAAAAAGTCAGAGAGAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT GCGTTCCTTCTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG AAAAACTCAGGCAGCACATTTCACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTCATG CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAACT AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAAATTTGACCAAAGTT TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGATGTCGCTGAGCTGGAACTCCAGA ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT AGTTTACAGAAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC CTTGGACCGCCTGCCAGCCCAGCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA AATATTCCCTTTGCAAATGGGATT**TAA**ACTAAGATAATATGCACAGTGATGTGCAGGAAC AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG ATACATCTTTTAAAATAAAACAGAGAGGGTGCATAGAAGGCTGATAGAAGACATAACTGAAT GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT TTCTTACTAAAAAAAAAAAAAAAA

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDFAFL
LHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKLRQHISRNAQDKQELHLFMLSG
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESLRELRHLKILHVKSNLTKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSLSNLQELDLKS
NNIRTIEEIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLESLPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP
EKVGQLSQLTQLELKGNCLDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP
FANGI

### Signal sequence:

amino acids 1-20

#### N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 326-330

#### Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

#### Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

### N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT GCTCAGTGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCGTGCTCTGTCTTTACTTCAAAATACACAAC GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACTGACGTTTCCCTGGAGGTGTCCAGAAA GCTGATGTAACACAGAGCCTATAAAAGCTGTCGGTCCTTAAGGCTGCCCAGCGCCTTGCCAAA**ATG**GAGCTTGTA AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCTGTTTTGGCGGAGGCTGACAATGGCGGAGGCTGAAGGCAAT GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC  ${\tt AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACCAAAAGCATCAGAATTATCTTTTCCTATGTCCAGCTT}$  ${\tt GATCCAGATGGAAGCTGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG}$  ${\tt CAAGTCTGCAGTAAAAACGACTATGTTCCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT}$ GGCGGTTACCTGGATACCTTGGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTTCAAAGAGATTTTCCTAGAAATAGAC AAACAGTGCAAATTTGATTTCTTGCCATCTATGATGGCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT GGCCGTGTGACTCCCACCTTCGAATCGTCATCAAACTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCT  ${\tt TACCGGGGATTTTCTGCTTCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT}$ GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA GACCCAACTTGCAGACCAAAATTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTTTGAATCCAATTCA  ${\tt TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACTCTTTTTGTTCAAGTTAGTCTGCAC}$ ACCTCAGATCCAAATTTGGTGGTGTTTTCTTGATACCTGTAGAGCCTCTCCCACCTCTGACTTTGCATCTCCAACC TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTAAGGTGTATCCCTTATTTGGACACTATGGGAGA TTCCAGTTTAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGAT AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG AAAACAGATTCCATCATAGGACCCATTCGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTCAGCATGAAACACATGCGGAAGAAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCCTTCATGGTTCTAGCTCTG AATGTGGTGACTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG  ${\tt CAGAACTAT} \textbf{TAA} {\tt CTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT}$ GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACGGCCTGCATGTAAAAAAA

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI ERPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSTLT FQIVTDSARIQRTVFVFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRP KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHA EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

### Signal sequence:

amino acids 1-24

#### Transmembrane domain:

amino acids 571-586

### N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374, 394-398, 419-423

# Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383, 408-412, 463-467, 520-524, 556-560

# Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

### N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

# Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

TGCTGCCGCCGCCGCCGTGCCCTGCCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGAT TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGC CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGGT CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCCAAGAGGGACATTGTCAAGGAA CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGACTGTACTATTCCCTTTTTGA ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG TCGGATGGTGACGGAGGACCCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA CATAAATGGGAAAACTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG GAGGAAATCTTTTGATGAATATTGGGCCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG GAGCGACTGAGGCAAGTGGGGTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA TACCTGGCGATCCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACATCAGGACAGCTGTTCCTTGGCCAT CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATC TAA AGTGCAGCAGAGTGGCTGATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA AAGCAATGTAAACTGGATAAGAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCCACTA AATTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT TTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG GTACAGTAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAATGTGAAGGCCTAGGACATTA GTTTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT TTTAAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA ATGTAA

MRPQELPRLAFPLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

### Signal sequence:

amino acids 1-28

### N-glycosylation site.

amino acids 171-175, 239-243, 377-381

#### Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

#### Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

### N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

#### Leucine zipper pattern.

amino acids 410-432

#### Alpha-L-fucosidase putative active site.

amino acids 283-295

TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC  ${\tt TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACG} {\tt ATG} {\tt GCCAGGTGCTTCAGCCTGGTGTTG}$ CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC AGCTGAATTTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACTCATCTGAT ACTTGGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA AACTGCAACACAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCTT ACTCTACAATACCTGCCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG AGAAAAAATTGATTTGTGTCACAGAAGTTTTTATGGAAACTAGCACCATGTCTACAGAAAC TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA  ${\tt TTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCCAAAGAACCAAAGAAGAAGAAGTCCACCCTT}$ GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC TTTCTAGCCTGGCTATGTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTATCTGA GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTCA GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTATTTTTTGCTGAGACTAATCTT ATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA GCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA AGCATTTAGAAAACTT

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR LLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQF AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFF GAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP SKTTVRCLEAEV

#### Signal sequence:

amino acids 1-16

#### Transmembrane domain:

amino acids 235-254

#### N-glycosylation site.

amino acids 53-57, 130-134, 289-293

#### Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

### Tyrosine kinase phosphorylation site.

amino acids 79-88

#### N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

 $\texttt{AG} \underline{\textbf{ATG}} \texttt{GCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCGCGACTTTCACGATGG}$ CTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG ACTGGAGAGAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC TGATGACGTGCAAACCCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA TTGGTCTAATGACTGCCAATCATTTGCCCCCTATCTATGCTGACCTCTCCCTTAAATACAACT  $\tt GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC$ AAAGTGAGCACATCACCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAAACTATCA AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTCAGA  ${\tt TGGGGAAAACAAGAAGGATAAA{ extbf{TAA}}{ extbf{GATCCTCACTTTGGCAGTGCTTCCTCTCTCTCAATT}}$  $\tt CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG$ CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG TCATAGGACGGAGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC TTCATTAAAAGTATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG TTTTTATCATTGAAGACAATATTGAACAACCCCCTATTTTGTGGGGATTGAGAAGGGGTGAA TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATT TCCTTTTG

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHGLPTQREDGNPCDFD WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

#### Signal sequence:

amino acids 1-48

#### Transmembrane domain:

amino acids 111-125

#### N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

#### Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

#### N-myristoylation site.

amino acids 188-194, 225-231

#### Myb DNA-binding domain repeat signature 1.

amino acids 244-253

GGACAGCTCGCGGCCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG GGGCCCAGCCTGGCCGGGTCACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTACTCCAACAG CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG TGGTGGAGACACCCGAGGAGACCCTGTTCACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT GTCGGAGAACGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG  ${\tt ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA}$ GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCA ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGGGGCCTGGACTGGTGCAACGCGGG  $\mathtt{CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC}$ CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCCTGCACCGCTATGAT GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCTGAGAAGCT GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT  ${\tt CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC}$ CGTTTTCCCTTGTGGGTTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT TTAAACATTTTTTACTATTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTGTGGGCCTTCTG GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTTCAGGAGGTTGGTGCCCAGAGTGGGC GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT  ${\sf CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCTTTGGGCCTCTCCTGCC}$ CACCTCTACTTCTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA 

MGLLLVPLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQ GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD KEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR HRRLHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

#### Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

GGAGAGCGGAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC CAGGGGAGGTGCACCAGGCGCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA CTTCCAGTACGACCATGAGGCTTTCCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA  $\tt CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC$ CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGAC GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGCGTTTCCGGGTGGC CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGGGGCTGACAGCCTTCCTGCACCCCGAGG AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA GGAGCCGGCGTGGGTGCAGACGGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCTGCCCAGGACCAGCCC CTGGTGGAAGCCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG  ${\tt ACCTGACCCGGCACCACGATGAGCTG} {\color{red}{\textbf{TGA}}} {\color{blue}{\textbf{GCACCGCGCACCTGCCACAGCCTCAGAGGCCCG}}$ GCAGATGCAGTCCCAGGCATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGGCTTC TGTCCCTGTCACACCCCCAACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATAC CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT GGCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA AAAAAAAAAAAAA

MMWRPSVLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT YDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE ELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ FRDFRDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV GSQATNYGEDLTRHHDEL

#### Signal sequence:

amino acids 1-20

#### N-glycosylation site.

amino acids 140-144

### Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293, 291-295, 298-302

#### N-myristoylation site.

amino acids 263-269, 311-317

### Endoplasmic reticulum targeting sequence.

amino acids 325-330

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG GCGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTCCCAG CCTGTCTGTCGTCGTTTTGGCGCCCCCGCCTCCCCGCGTGCGGGGTTGCACACCGATCCTG GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG CGGGTCGTCTGTGTCCTCTCTCTCTGCGCCGCGCCCGGGGATCCGAAGGGTGCGGGGCTCT GAGGAGGTGACGCGGGGCCTCCCGCACCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG GTGTGAGCAGCCTATCAGTCACCATGTCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG TGTCTGCTGCTGCCGGGGCCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATG TTTTACCAGAGGCTTGGACATCAGGAAAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC TGGTCGAGAAAACTATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT GGTCTGCTTCTTTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACTGG GCCGATTTAATTTACAGAAGAATTTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA GAAGGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAA AAACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGGTA ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA GTAAGAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAA GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCCTC ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT GATGCCATTTCCTTCACTGTTAGAAATGTGTTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA CTTCCTAGTAATTGTCACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCACAAGAGAGTTCACAGGATT  ${\tt AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAA} {\color{red} \textbf{\underline{T}}}$ **AA**TGGTAACATTTTGACAACTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACTATTAAGTATGTCAAC AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA GATAATGTGGATTAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA 

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFSVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI
FSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

#### Signal sequence:

amino acids 1-24

#### N-glycosylation site.

amino acids 100-104, 221-225

#### Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424, 425-429, 478-482, 528-532

#### N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211, 239-245, 240-246, 261-267, 403-409, 442-448, 443-449

#### Amidation site.

amino acids 145-149

CGCCGCGCTCCCGCACCCGCGCCCCACCGCGCCGCTCCCGCATCTGCACCCGCAGCCC GGGCGGCGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC TGCTGGCGGCGGTCCCACGGCCCCGCGCCCGCTCCGACGGCGACCTCGGCTCCAGTC AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA GGTTGAGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG CAGAAGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCAGCTAT TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG TGGGAGACGAAGAAGGCAGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG GTGTATGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG AGGACCTGGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT GCACTGCTGGGAGGGGAAGAGATT**TAG**ATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA TCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGC TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG ACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG AGTCTCCCTCTGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAA CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCCTGCATTACATGTGTTTATTCATCC AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG GCTCAGAGACTGCAAGCTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT CTGGTTGTGACTCTAAGCTCAGTGCTCTCCCACTACCCCACACCAGCCTTGGTGCCACCAA AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG GTCAAACTAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT TGCATTAGTAACTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCA TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA TAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC 

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

#### Signal sequence:

amino acids 1-19

#### N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

### Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

#### N-myristoylation site.

amino acids 202-208, 217-223

#### Amidation site.

amino acids 140-144

TAAAGCCTTAAGGACAGGCCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTTGAGTCAAAAAAACATGGGAGGG CCAGGCACGGTGACTCACACCTGTAATCCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTCAGGAG TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC AGGTGCCTGTAATCCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGTGGGGTAGATACTGCTTCTCTGCAACCTCCTTAACTCTGCATCCTCTTCTTCCAGGGCTGCCCCTGATGGGGCCTG GTGTAGAATGACTGCCCTGGGAGGGTGGTTCCTTGGGCCCTGGCAGGGTTGCTGACCCTTACCCTGCAAAACACA ACTCTTGCTAGCTTGGGTGGCTGGTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCTCA GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCCGCTCGTCCTACCGCGAGGCTACCACTGTGGACTGCAATGA  $\tt CCTATTCCTGACGGCAGTCCCCCGGCACTCCCCGCAGGCACACAGACCCTGCTCCTGCAGAGCAACAGCATTGT$  $\tt CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCCAGAACAGCTTTTCGGA$ TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT GGAGGACCACAGCTTTGCAGGGCTGGCCAGCCTACAGGAACTCTATCTCAACCACAACCAGCTCTACCGCATCGC GAACTTCCGGCCCCTGGCCAACCTGCGTAGCCTGGTGCTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC  $\tt CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGCTGGCCCGGGTGCCCAGGCGGGCACT$ GGAACAGGTGCCCGGGCTCAAGTTCCTAGACCTCAACAAGAACCCGCTCCAGCGGGTAGGGCCGGGGACTTTGC GAACCTCCCGAGCTGACCAAGCTGGACATCACCAATAACCCACGGCTGTCCTTCATCCACCCCCGCGCCTTCCA CCACCTGCCCCAGATGGAGACCCTCATGCTCAACAACAACGCTCTCAGTGCCTTGCACCAGCAGACGGTGGAGTC GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGCCCCTCATCTCCCCACGAAGCTTCCCCCCAAG  $\tt CCTCCAGGTAGCCAGTGGAGAGAGCATGGTGCTGCATTGCCGGGCACTGGCCGAACCCGAACCCGAGATCTACTG$ GGAGCTGCGGAGGGTGACAGCAGAAGAGGCAGGGCTATACACCTGTGTGGCCCAGAACCTGGTGGGGGCTGACAC  $\tt GTTGGCTTGTGTATGGGCCAGGACCAAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCCTGGGCTCAT$ TGCCATCCTGGCTCTCGCTGTCCTTCTCCTGGCAGCTGGGCTAGCGGCCCACCTTGGCACAGGCCAACCCAGGAA GGGTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTTCTGGGGGCTGGAGTGCCCCTTCTGTCCGGGTTGT GTCTGCTCCCTCGTCCTGCCCTGGAATCCAGGGAGGAAGCTGCCCAGATCCTCAGAAGGGGAGACACTGTTGCC  ${\tt ACCATTGTCTCAAAATTCT} \underline{{\tt TGA}} {\tt AGCTCAGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTACCAAA$ A GAGAAG CAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCTGGCAGCTGGGC $\tt CCTCTGCTGCCATCTGAGGAACATCTCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCATCTT$ CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA AAAA

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA VPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE MLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLP ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIR CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREMTDHCLPLISPRSFPPSLQ VASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAEEAGLYT CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS CHRALGDRPGLIAILALAVLLLAAGLAAHLGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRVV SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

#### Signal sequence:

amino acids 1-18

#### Transmembrane domain:

amino acids 629-648

#### N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 485-489

#### Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146, 243-247, 313-317, 488-492, 700-704

#### Tyrosine kinase phosphorylation site.

amino acids 532-540

#### N-myristoylation site.

amino acids 15-21, 493-499, 566-572

#### Amidation site.

amino acids 470-474, 660-664, 692-696

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT  $\mathtt{CTAC} \underline{\mathtt{ATG}} \mathtt{CGGCAGACCATCAAGGTGATCAAGGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA}$ CATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCCACCC  ${\tt CCTGGCCACACTCTTCAAGATCCTGGCGTCCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA}$ CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA GCGCTTCGCCGTCTTCCTGTCGGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT GGACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTCATGCTCAGTGGCAT CAGCATTGCCCAGCTCACGGGCCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT GGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCCTCAACAGCCTCAA GAAGATGGCGAACCTGACTGGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCCACTCCATCTTCAGCCT CCACAACCTGCAGGAGATTGACCTCAAGGACAACCATCAAGACCATCGAGGAGATCATCAGCTTCCAGCACCT GCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA CTACCTGGACCTCAGCCACAACAACCTGACCTTCCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCCTGAGCGAG GCCGGCCCAGCACAGCAGCAGGACCGCTGCCCAGTCCTCAGGCCCGGAGGGGCAGGCCTAGCTTCTCCCAG AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCAGGAGCCTGGGGCCGCTTGTGAGTCAGGCCAGAGCGAGA  ${\tt AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAGCTCTCCCCCAGGGGCTGAGCTCTCCCCCAGGGGCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCTGAGCTCAGGGCCCAGCTCTCCCCCCAGGGGCCTGAGCTCTCCCCCAGGGGCCTGAGCTCTCCCCCCAGGGGCCCAGCTCTCCCCCCAGGGGCCCAGCTCTCCCCCAGGGCCCAGCTCTCCCCCAGGGCCCAGCTCTCCCCCCAGGGGCCCAGCTCTCCCCCCAGGGGCCCAGCCTCCAGCTCCAGCTCCCCCAGGGCCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCAGCTCCCCCAGGGCCCAGCTCCAGCTCCAGCTCCAGCTCAGCTCCAGCTCCAGCTCAGCTCCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGCTCAGC$  ${\tt AGATAACTTATACATTCCCAAGAAAGTTCAGCCCAGATGGAAGGTGTTCAGGGGAAAGGTGGGCTTTTCCCCC}$ CCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCTCCAGCTGGA AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTCCCTTGAGCAAAGCAGCCAGACGT TGAACTGTGTTTCCTTGGGCGCAGGGTGCAGGTGTCTTCCGGATCTGGTGACCTTGGTCCAGGAGTTATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTCGGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGAGGAG CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTCCCCACCTCCTGCGGCATGGGTGTGT CCAGTGCCACCGCTGGCTCCGCTGCTTCCATCAGCCCTGTCGCCACCTGGTCCTTCATGAAGAGCAGACACTTA GAGGCTGGTCGGGAGTCGCCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCTGGCGC  $\tt CTGGAGTGCACAGCCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT$ A GAAGGGTCCCCGCCTTAGATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT ${\tt CCATCCGTCTGTCCATTTGTGTTTTTCTGCGTCGTGTCATTGGATATAATCCTCAGAAATAATGCACACTAG}$ CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACTTGTAGACTCGGTCACAGTATCAAATAAA ATCTATAACAGAAAAAAAAAAAAA

MRQTIIKVIKFILIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK RFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTDIKEIPLWI YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIISFQ HLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLT FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQSLPSRVGELTNLTQIE LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

#### Transmembrane domain:

amino acids 51-75 (type II)

#### N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 85-89

#### Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368, 398-402, 493-497

#### N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

 $\texttt{GCCTGTTGCTGATGCTGCGGTGCGGTACTTGTC} \textbf{AT} \underline{\textbf{G}} \textbf{GAGCTGGCACTGCGGCGCTCTCCCGT}$  $\tt CCCGCGGTGGTTGCTGCTGCTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT$ GGCCCACAGAGGAGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCAT GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAAACTTTGAGGAAATTGGGC CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT GTGGATAATCCCGTGGGCACTGGGTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC ATTGGTCTAGAGCTTTATAAGGCCATTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT TGCCTTGGGTGATTCCTGGATCTCCCCTGTTGATTCGGTGCTCTCCTGGGGACCTTACCTGT ACAGCATGTCTCTCCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT GATCATTGAACAGAACACAGATGGGGTGAACTTCTATAACATCTTAACTAAAAGCACTCCCA CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAGGCCTG ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT  ${\tt GATGAGACTGGTGACTCAGCAAGAA} {\color{red}{\textbf{TAG}}} {\tt GATGGGTGGGGCTGGAGATGAGCTGGTTTGGCCT}$ TGGGGCACAGAGCTGAGCCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAAACCTAAGATTTTTTAAAAAAATTGAT TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNKGLYRE ATELWGKAEMIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDALS QLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL IVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP SDQGDMALKMMRLVTQQE

#### Signal sequence:

amino acids 1-25

### N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 101-105

### Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

### N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175, 187-193, 195-201, 331-337, 332-338, 360-366

GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGGCCGTTATCAGGACCATGCGGCCGA  $\tt CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCA$ GGGGAGCCTGCGCGTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG CACTCACGGCGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG  ${\tt TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG}$  ${\tt CCTTGGTGAAGCTGTCTGCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG}$  ${\tt GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGGGGGTACATCAA}$ AGAGGATGAGGCACTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG  $\tt GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT$  $\tt GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG$ GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG  $\tt CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTCCCTCT$  ${\tt TCTCTGGGCTCTCCCACTCCTGGGGCCGGTC} \underline{{\tt TGA}} {\tt GCCTACCTGAGCCCATGCAGCCTGGGGCC}$ CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTTGGTAATAAACACATTCCAGTTGA 

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP SPHTLQEVQVAIINNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDSGGPLACNKNG LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL LGPV

#### Signal sequence:

amino acids 1-18

#### N-glycosylation site.

amino acids 167-171, 200-204, 273-277

### Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

#### N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245, 259-265, 269-275

#### Amidation site.

amino acids 33-37

### Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 252-263,

Serine proteases, trypsin family, histidine active site.

amino acids 78-84

 $\tt CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGA\underline{ATG} \tt GGACTCCAAGCCTGCCTCCTAGGGCT$ CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC  $\tt TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT$ GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAG CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCT GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTCCC TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCCATGTGGACTTTGTGGGGGGGACTGCACCGT TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT GCATCTGGGGGTAACCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGAC TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCCGGGATTGAGGCCAGTCTAGATGTGCAGT ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG GGACAGGAGCCCTTCCTGCAGTGGCTCATGCTCAGTAATGAGTCAGCCCTGCCACATGT GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGACAGT GGGGCCGGGTGTTGGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG CCCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAA TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCCACGGCCTTCATACCAGGAG GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCAATGC CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTCAGCAACA GAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTTGGGGGGGATCCTA GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC  ${\tt TGGGGAACACCCAACTTCCCAGCTTTGC} \underline{{\tt TGA}} {\tt AGACTCTACTCAACCCCTGACCCTTTCCTATC}$ AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG GAAGCCCTGCTGAACCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA TGCTGTGAGCTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA TTTCACTTGATATTCATTCCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTTACTCT TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG GCCTTTCCATCATAGTTGCCCACTCCCTCTCTTACTTAGCTTCCAGGTCTTAACTTCTCTG ACTACTCTTGTCTTCCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA TGTAAAAA

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGOGFCSGPGWDPVTGWGTPTSOLC

### Signal sequence:

amino acids 1-16

#### N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

#### Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

#### Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

#### N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225, 248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488, 521-527, 533-539, 549-555

GCCGCGCGCTCTCTCCCGGCGCCCACACCTGTCTGAGCGGCGCAGCGAGCCGCGGCCCGGGC GGGCTGCTCGGCGGAACAGTGCTCGGCATGCAGGGATTCCAGGGCTCCTCTTCCTC TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCCCTGGAAACCCACTTG GCCTGCATACCGCCTCCCTGTCGTCTTGCCCCAGTCTACCCTCAATTTAGCCAAGCCAGACT TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCCAGTGTCATAAGGGAACT CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG CAGCCGCACAGAGACGCAGGTGGGCCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAAC ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGC AGGTTCAGCATTTTTGGGAAGGACTTCCTGCTCAACTACCCTTTCTCAACATCAGTGAAGTT ATCCACGGGCTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCCTCACAGCTGCCACTGCA TACACGATGGAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC AAGTTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA ATTTCAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCCACAAGAGAAAA TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATTCACTTCTC TGGTTATGACAATGACCGACCAGGCAATTTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA CCTATGACTTGCTCTACCAGCAATGCGATGCCCAGCCAGGGGCCAGCGGGTCTGGGGTCTAT GTGAGGATGTGGAAGAGACAGCAGAAGTGGGAGCGAAAAATTATTGGCATTTTTTCAGG GCACCAGTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTCAGAATCACTC  ${f TGA}$ CACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTTCTTATTTTAGGAGAGGCC CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAACTG GTTTGTGTATCATATCATATCATTTAAGCAGTTTGAAGGCATACTTTTGCATAGAAATAA AAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG CAAACTTTGATTTTATTTCATCTGAACTTGTTTCAAAGATTTATATTAAATATTTGGCATA CAAGAGATATGAAAAAAAAAAAAAAA

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD AQPGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW IKGNYLDCREG

#### Signal sequence:

amino acids 1-19

#### N-glycosylation site.

amino acids 93-97, 207-211

### Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

### Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

### N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCTGCTCCCCCGCCCCACCAGCCATGGTGGTTT CTGGAGCGCCCCAGCCCTGGGTGGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTGGG CCCACCTGTGTATTCCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCGTCTCGAG CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC  $\tt CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCCTATCATCGACTCGGAAGTCTGCAGCCAT$ CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT GGAGGGGAGCGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG GCGCCTGGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA GCTCCGCGGGCGCTCAGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCCAC GACTTCCGGCCCCGCCCCCGGGCCCCAGCGCTTTTGTGTATAAAATGTTAATGATTTTTAT CTCCAAAAAAAAAA

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318</pre>

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG SGAAARS

### Signal sequence:

amino acids 1-32

### N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

#### Glycosaminoglycan attachment site.

amino acids 826-830

#### Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

#### Tyrosine kinase phosphorylation site.

amino acids 607-615

#### N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

GACGGCTGGCCACC<u>ATG</u>CACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA GATGGGACGAGGGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCAC AACAAGGAGCGCGGCGCGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGT  ${\tt GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT}$ GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT  ${\tt TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGTGGAAACCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCCAGGCCCCAGGCCCAGGCCCAGGCCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCCAGGCCCCAGGCCCCAGGCCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCAGGCCCAGGCCAGGCCAGGCCCAGGCCAGGCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCAGGCCAGGCCAGGCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCAGGCCCCAGGCCCAGGCCCAGGCCCCAGGCCCAGGCCCAGGCCCAGC$  ${\tt CCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC}$ AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG GGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA TGCCACGGGTGGCCTGGCTCTGCAGTCGTCCTTGCCAGGTGCAGAGGGCCCTGACA  ${f A}{f G}{f C}{f C}{f T}{f G}{f G}{f G}{f C}{f C}{f$  $\tt CTACTGCTCCTCCTCTGGTGTTTGGCTGGAATCTTC\underline{TGA}{ATGGGATACCACTCAAAGGG}$ TGAAGAGGTCAGCTGTCCTGTCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATA  $\tt CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC$ ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACC GGGCCCACACCTCTCCTGCCCCTCCTGAGTCCTGGGGGGTGGGAGGATTTGAGGGAGCT CACTGCCTACCTGGGCCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG TGTAGCTGGGGATGGGGATTCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC TTTGAGTGGGGGGGGGGGGGGGGGGGGAGGGAAGTAACTCCTGACTCTCCAATAAAACCT GTCCAACCTGTGAAA

MHGSCSFLMLLLPLLLLVATTGPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEE LAAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ MCGHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELLVCNYEPPGNVKGKRPYQEGTPCSQC PSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLVTEV SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAELPPSSEVLAS VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV SGLNSGPGHVWGPLLGLLLLPPLVLAGIF

#### Signal sequence:

amino acids 1-22

#### N-glycosylation site.

amino acids 114-118, 403-407, 409-413

#### Glycosaminoglycan attachment site.

amino acids 439-443

#### Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

#### N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

#### Amidation site.

amino acids 82-86, 172-176

#### Peroxidases proximal heme-ligand signature.

amino acids 287-298

#### Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

#### Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

 $\tt GTAACTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAATTCGGTCATTCTCCAAGTT\underline{ATG} GTGGACGT$ ACTTCTGTTGTTCTCCCTCTGCTTGCTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTCAT ACATCTGAAAGAGTTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATT TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATAC ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAACAATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAG GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA TCTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATT  ${\tt TCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGA}$ TGGCTTTGTGTGTGATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTC CAATTTGAGTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCCAATGACTTTTGCTTGGAAAAAAAGACAATGA ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCAT  $\tt CCGAGCTGGGGCCATGGCACGCTTGGAGTGTGCTGTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGA$  $\tt GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGC$ AACTCTGACTGTCCTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGC  $\tt CGTCCTACAGTGCATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGACCAAAGATGATAGCCCATTGGTGGT$ AACCGAGAGGCACTTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAA  $\tt CTGCGACTCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGC$ TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTT AGCTGACAGGCAGGATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGG CACAGATCTGTTCCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA TCCTTTTGAAACATATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTA CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATC GTGGCCTTCACATGTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTG TCTAAACAAGTCCTCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGG TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCC AGATTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTC GAACCAAAATTACAAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTA TTTTTTAACTTTGTTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTT  ${\tt TTATAATGCCAGATTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTT}$ TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP GYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKM QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE FCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE ISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLOGNAFSO MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCD DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQG GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA MARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN SAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTER HFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDG WATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLAD RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP MYLKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDA YSSFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQS YDLDT

#### Signal sequence:

amino acids 1-19

#### Transmembrane domain:

amino acids 746-765

#### N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

#### Glycosaminoglycan attachment site.

amino acids 826-830

### Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

### Tyrosine kinase phosphorylation site.

amino acids 607-615

#### N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGAAGAACAAC GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA  $\tt CCACCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTTTTCTC$ TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT GGTGTGGTGTTTTCCTTTCTTTTTGAATTTCCCACAAGAGGAGAGAAATTAATAATACATCTGCAAAGAAA TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCTGGATGTTGCT  $\tt GTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA{\color{blue}ATG}TTGAACAAGAT$ GACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCTGCTTGTGGTGCT GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA  ${\tt CCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACGGCT}$ GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT GCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGG GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACTTGAGGTATTTGAACCTTGC CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAGCTGGATCTTTCTGGGAATCA TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACTGTGGATGATACAGTCCCAGAT TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTTGGCACCAATAATCTAAC ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACTG TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT GATTGTGGAGCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAAGTGTCGGGCCTCCACATC  ${\tt CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCGGATAGCTGT}$ GCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA TTCCGTTGGGAATACTACTGCTTCAGCCACCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC AGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTCGACAGAGAAAACCTT CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTCATGAAGACTACCAAAATCATCAT TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA TCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC CATGGAAAGCCACCTGCCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT CAACCACAACAACAGTTAACACAATAAATTCAATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAAGTGTCTTTACAAAAAAACAA AAAAGAAAAGAAATTTATTTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN
TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

#### Signal sequence:

amino acids 1-44

#### Transmembrane domain:

amino acids 523-543

#### N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 183-187

### Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

#### N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

 ${\tt AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGGCA}$ GAGGGCGGGCGTGCACCCTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCCGCCCGGCTGGGAGCTTCGGGTAGA GCGGTGCTGGGCGCTGGCCGGTCCGACAGCGGCGGTCGCGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC GAGCGCCCATGCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGCAGTCGTAAGCGGCTAGCGCGTCTT  $\tt CCCGAGCCACTCCCGTCCTGGGTCGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC$ ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA GTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG TTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATTTCCAGCCCTACAG CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA  $\hbox{\tt CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT}$ CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG AATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAGGCGACTGATACTC CAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATTGCATTTAAATACA GATGATTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGAACTACTGCATGAT  $\tt CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTC$ AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC TTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA ATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACCGAGAGGCAC TTTTTTGCAGCAGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG ATGTCTAACACCCTTGGCACTGAGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCCT  ${\tt CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGT}$  ${ t GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC}$ AACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTTAGCTGACAGGCAG  ${\tt GATGGGTACGTGTCTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGGATTTTTCTTACCA}$  ${\tt CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTC}$  $\tt CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGATCCTTTTGAAACA$ TATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG  ${\tt GAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT}$ GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC TCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCCAAGAGCCTTTTAT  ${ t GAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC$ CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAAGGATAAAAATGCTTTATTATACAGATGAACCAAAATTAC AAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTAACTTTG  ${\tt TTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTATAATGCCAGA}$  ${\tt TTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAATAGAAGTT}$ 

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY FDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKMOR NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK KLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQGGE VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA RLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA TVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLADRQ DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY LKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS  ${\tt HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDAYS}$ SFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

### Signal sequence:

amino acids 1-27

#### Transmembrane domain:

amino acids 808-828

### N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519, 688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

### Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378, 383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735, 799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022, 1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

### N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433, 513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

CAAAACTTGCGTCGCGGAGAGCGCCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCCCAGCTGAGAC TGGGGGAGCGCTTCGGCCTGTGGGGCGCCGCTCGGCGCGCGGGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCCTGGGTGGTCCCGTCCCCTATCCCTCTTTATATA GAAACCTTCCACACTGGGAAGGCAGCGGCGAGGCAGGAGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG TGCTCCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGG AGCCCGGGCCTGGCCCAGCCGCGGTCAGCTGCCCCCGAGACTGTGCCTGTTCCCAGGAGGGCGTCGTGGACTGTG TGGAAAAGATCTACCCTGAGGAGCTCTCCCGGCTGCACCGGCTGGAGACCTGCAAAACAACCGCCTGA  $\tt CTTCCCGAGGGCTCCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC$ TGACCTTGGCACCCCGCTTCCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT ATGGGCTCACCTTTGGCCAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC TGCCGGACAACATGTTCAACGGCTCCAGCAACGTCGAGGTCCTCATCCTGTCCAGCAACTTCCTGCGCCACGTGC  ${\tt CCAAGCACCTGCCGCCTGTACAAGCTGCACCTCAAGAACAAGCTGGAGAAGATCCCCCCGGGGGGCCT}$ CGCGCAGCCTGGTGCTGCACCTTGGAGAAGAACGCCATCCGGAGCGTGGACGCGAATGTGCTGACCCCCATCC GCACCCTCATGATCCTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCCTGGAGG AGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGGTGCACCGCGACGCCTTCCGCAAGCTGCGCCTGCTGC GCTCGCTGGACCTGTCGGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAAGG TCAAGCGCAATGAGCTGGCTGGCCTGGCAGGGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA  ${\tt CCAGCAACCGACCGAGCCGAGCCCTGGGCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA}$ TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAACA AGATTAGTGCGGTGCCCCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAACAAGC TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT  ${\tt AGGAAGAGAACAAGATAGTGACAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGGACTCTTTTCTGC}$ AGCACACGCCTGTGTGCTGTGAGCCCCCCACTCTGCCGTGCTCACACAGACACACCCCAGCTGCACACATGAGGCA TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCCACGGCGTGTCCCACGGCCAGACACATGC GGAACTCACAAAAGCTGGCTTTTATTCCTTTCCCATCCTATGGGGACAGGAGCCTTCAGGACTGCTGGCCTGGCC 

MEGEEAEQPAWFHQPWRPGASDSAPPAGTMAQSRVLLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLHRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEEE

#### Signal sequence:

amino acids 1-48

#### N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

#### Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

#### N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341, 477-483, 498-502, 539-545, 548-554

#### Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493, 535-557

GGGAGGGGGCTCCGGGCGCGCGCGCAGCAGACCTGCTCCGGCCGCGCGCCCCCGCTGTCCTCCGGGAGCGGCAG CAGTAGCCCGGGCGAGGGCTGGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCACCACCC GCTGGTATCCTCGCTCTACCTGCAGGCGGCCGCCGAGTTCGACGGGAGGTGGCCCAGGCAAATAGTGTCATCGAT TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTCAGCC TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTCATCCTGGTTATGCTGG AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA CTCCATGGCAAACTGTCAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCATAAAGGCTTCGATCTCATGTATATTGGAGG AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA TAATAATTGGATTCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCCAGAAAGGCCCAACCACCGGACTGACAACTATAGC TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCGGCAGCCAAAGCCCCAGG GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTCAGGGGACCTGTGCCTGTCATTCAGGCA CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCCACGGAGCAGCCCTGTG GGGAAGAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA  $\texttt{AAGA} \underline{\textbf{TGA}} \texttt{TTAAAGGGTTGGAAAAAAAGATCTATGATGGAAAATTAAAGGAACTGGGATTATTGAGCCTGGAGAAG}$ AGAAGACTGAGGGGCAAACCATTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTTGGCGACCAGCTGTTCTCCATATGCACTAAGAATAGAACAAGAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGCAGG TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG CATTTCCCTCTTGACCTCCTAATGGAGAGGGGATTGAAAGGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC ATTTTCATCGGGTGCATTCTCTCTGCTTCGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC ACACCGGCAGACCTTTCCTTCACCTCATCAGTATGATTCAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC AGAACAGTAATATTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCATTTAACCTGGTAAAGGCAGGGCTGG AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTCATT  ${\tt TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTT}$ 

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQPRCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLTCVYIPKVMIEPSGPIHVPKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDNRVQTDPQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRLMHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSESQR

### Signal sequence:

amino acids 1-17

#### N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

### N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242, 421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCCTTTAGATTGTGA CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCA GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT  $\tt CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG$ CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTAT GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACT  ${\tt CCTGGATTAGGAGGACTGACAATACTACATATCATTAAGCATGGGCCTCGCTTAGAAGTT}$ GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC CGGCAGGCAAGATGAAACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG CACAGAAAGGAAAATCATTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATT ATATCCATGTGTCTTCTCTTCCTATGGAAAAATATCAACCCTACAAAGTTATAAAACAGAA ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT GTATGAAGTTATTCAGCACATCCCTGCCCAGCAGCAAGACCATCCAGAG**TGA**ACTTTCATGG GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT GTGCAACAGAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA TGTGGAAACTTTACATTGTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA CAAAGTAATAAGGATGGTTGTCACAAAAACAAAACTATGCCTTCTCTTTTTTTCAATCACC AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA 

MWLKVFTTFLSFATGACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSVPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

### Signal sequence:

amino acids 1-18

### Transmembrane domain:

amino acids 341-359

### N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

### Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

#### Tyrosine kinase phosphorylation site.

amino acids 272-280

#### N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

### Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

 $\tt GCAAGCGGCGAA{\color{red} \underline{\textbf{ATG}}} GCGCCCTCCGGGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTT$ GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTCGCGTCATCACGGACGAGA ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATG TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC AGATTGCCTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCTTCAAAAAAAT TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAG AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTA CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG TTTCACTGTGTGAAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTGTCTCAAG AAATGTGTATTTCAGTGACAATTTCGTGGTCTTTTTAGAGGTATATTCCAAAATTTCCTTGT TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC GTTTCAAACTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAATT TTGGCCACTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA TTTTACATTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG CATCTTCTTGTATATGTCTTAAATGTATTTTTGTCCTCATATACAGAAAGTTCTTAATTGAT TTTACAGTCTGTAATGCTTGATGTTTTAAAATAATAACATTTTTATATTTTTTAAAAGACAA ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTCACAGGTAAAA GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT CACCTTAAAAGAAGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGAC AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGAATCCATTAGATTTACAG TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAAATATTGTACCATT AAAGAGTTTGGATGTGAACTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT 

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL QPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL ATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEEQEADEEDVSE EEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

#### Signal sequence:

amino acids 1-26

#### Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTT CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATTGAAAACTCTGGGTGTCTGC ATTGCTGATGGCCTGGTTTGGTGTCCTGAGCTGTGCAGGCCGAATTCTTCACCTCTATTG GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCCTGTGAATGCCTACAAACTGG TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCCTGCAGGACTCAGCTGCA GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA GAGGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA GGAAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAGAAGGCA TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCGTGGG GAGGGTGTCAAACTGACACCCCGTAGACAGAAGAGGGCTTTTCTGTAGGTACCACCATGGCAA CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGGACGAGTGGGACAGCCCGCACA TCGTCAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTA CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT  ${\tt TGAGACCTTGTGGATCAACAGAAGTTGAC}{\tt TGACATCCTTTTCTGTCCTTCCCCTTCCTGGTC$ CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCCTTTGTATGTTCCTATCAGGCT GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT GTGACTGAAGTCCCAGCCCTTCCATTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTTTTTAT 

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIANLSVQRQFFP TDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDYYHTV LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLR YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTPRRQKRLF CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYYDVMSDEEIERIKEIAKPKLARATVRDPKTG VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR TRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEVD

#### Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270, 346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

AGACAGGACAATCTTCTTGGGGATGCTGGTCCTGGAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC  ${ t ATCCCTGTGTCGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTCGAGAGCTCGGCTAGACCAAAGTG}$ ATGAAGACTTCAAACCCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGCTGGTGGCTGTCCTGACCTCCCGAGCTACACTGTCCA CCCGGGCTCCAGCAGGATGCAGGTGTGTCTCATGGGGATGAGCGGCCCGCCTGGCTCATGTCAGAGACCCTGC GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC  $\tt CCCGCCTGGCAGCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCATTG$ GCGCAGGCGAGCAGGCCCGGTACTGTCATGGGGGGCTTTGGCTACCTGTTGTCACGGAGTCTCCTGCTTCGTCTGC GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCCTGACGAGTGGCTTGGACGCTGCCTCATTG ACTCTCTGGGCGTCGGCTGTGTCTCACAGCACCAGGGGCAGCAGTATCGCTCATTTGAACTGGCCAAAAATAGGG ACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCTTCGCCGTGCACCCTGTCTCCGAAGGTACCCTCATGT ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA TCCGGAACCTGACCGTGCTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCCTGCTCCTTTCA CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACCCTTCTCCTGTGCAGATGGGG CTCCCAAGTGCCCACTACAGGGGGCTAGCAGGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA ATCGGCGCTATCAGCCCCGCCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC TGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCGAAGTCCTCAACCGCTGTCGCATGAATGCCATCTCTGGCT GGCAGGCCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCACCACAGAGATCACCCCCAGGGCCCC ACCTCTTTCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGGCTAGGGGGCCGTGCCCAGCTGGCTATGGCTCTCTTTG AGCAGGAGCCAATAGCACT TAGCCCCCCCGGGGGGCCCTAACCTCATTACCTTTCCTTTGTCTGCCTCAGCC ACATGTCTTCTGCC

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLD QSDEDFKPRIVPYYRDPNKPYKKVLRTRYIQTELGSRERLLVAVLTSRATLSTLAVAVNRTV AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV LGWDYFTEQHTFSCADGAPKCPLQGASRADVGDALETALEQLNRRYQPRLRFQKQRLLNGYR RFDPARGMEYTLDLLLECVTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLPLL VAEAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDPFLGVKAAAAELERRYPG TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFFP VHFQEFNPALSPQRSPPGPPGAGPDPPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA RARLAGELAGQEEEEALEGLEVMDVFLRFSGLHLFRAVEPGLVQKFSLRDCSPRLSEELYHR CRLSNLEGLGGRAQLAMALFEQEQANST

#### Signal sequence:

amino acids 1-15

#### Transmembrane domain:

amino acids 489-507

#### N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

#### N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGCTGCGCTTTCCTGTCCCCAAGCC GTTCTAGACGCGGGAAAA**ATG**CTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAGAAGATATCTTGAAAATTTCA GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAAAACC CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACTG GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTTAA AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCCAATGACTTATCACCC ATCAGATGCATGTGATGTATGGGGTATACCGCCTTAGGGCATTTTGGGCATATTTTCAAT GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG TGAATATGATCTTTGTATAGGACGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAA TACAGCTGTATGTTTCTTTTTCTTAATTTGGTGGCACTGGTATAACCACACATTAAAG TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTTCTTTAAAACACATGAACATTGTAAATG GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT TGTGATTAAAGTAAAACTTTTAGCTGTGTTTCCCTTTACTTCTAATACTGATTTATGTTCT AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACTAAGCAACTAAAGAAA ATTAAAGTGAAAGTTGAAAAAT

MLSESSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK AYKYAFDKYRDQYNWFFLARPTTFAIIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL PPNGSDND

#### Signal sequence:

amino acids 1-33

### N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

#### N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

 $\tt CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT$ GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTCGCAT CTTTCAAACCTGTGATGAGAAAAGTTTCAGCTACCTGAGAATTTCACAGAGCTCTCCTGCT ACAATTATGGATCAGGTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC AGCTGCTACTTCTTTTCTACTGACACCATTTCCTGGGCGTTAAGTTTAAAGAACTGCTCAGC CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA AACCTAAAATGAGAGAGTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA TGGGTGGACGCCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG  ${\tt AACAAAGGAAAATCTCTT} \underline{{\tt TAA}} {\tt GAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA}$ AGAACATGGCCACCCCCCCCCCCCCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC AAAAA

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN FTELSCYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

### Signal sequence:

amino acids 1-42

### N-glycosylation site.

amino acids 2-6, 62-66, 107-111

### Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

### N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

TCCGGCTCTGCGCTGCCTGACTTCTTCCTGCTGCTTCTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC TCAAATCCAGCAATCGAACCCCAGTGGTACAGGAATTTGAAAGTGTGGAACTGTCTTGCATCATTACGGATTCGC AGACAAGTGACCCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTGACAACAAAA TTCAGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA  $\tt CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC$ ACTGCCAGGAGTGAGGGCCACCCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCCACGGATT CCAGAGCCAATCCCAGATTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCACTGCTG TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA  ${\tt TGATC} \underline{{\tt TGA}} {\tt GACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA}$ GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTCAGAAGCTTTTCGTTTTTGGCCAAAGTTGACCA CTACTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTCTGGCCTGATTCCCGCATGAGTATTAGG GTGATCTTAAAGAGTTTGCTCACGTAAACGCCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCACCACTGGTCGTT CAGCAGCACGACACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGGCGGGAACCCA GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCCTGCTTGTCCAACAGGGTGTCAGGATTTAAGGAAA TACATCTAAATTTTTGCTAAGGATGTATTTTGATTATTGAAAAGAAAATTTCTATTTAAACTGTAAATATATTGT CATACAATGTTAAATAACCTATTTTTTAAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTTAAAT TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT CACACAGTTTTAGCCTTTTTCACAAGGGAACTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT TAAAAATTCCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT  ${\tt TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT}$ AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTAAAAGAAAATGGAT GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAGGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG TGAAACGCCTGAATCAAAAGCAGTTTTCTAATTTTGACTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCATT TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGA TCTTGGTTGTCATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA TGGGAACCAGGTCTGAAAAAGTAGAGAGAGAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA CGGAAAAGGAATACTCGTGTATTTTAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT  ${\tt GCCTTTGGATGGATGTTGCTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAC}$ CTCATTTATAAAAGCTTCAAAAAAACCCA

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSD PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNDRK EIDEIVIELTVQVKPVTPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRA NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGG VLVVLAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267